Ravnomjerna raspodjela broja birača po izbornim jedinicama temeljem matematičkog modela

Kristian Sabo[*]
Rudolf Scitovski[**]
Petar Taler[***]

UDK 342.828
Izvorni znanstveni rad / original scientific paper
Primljeno / received: 29. 10. 2011.
Prihvaćeno / accepted: 9. 2. 2012.

U radu je predložen matematički model na osnovi kojeg je moguće definirati maksimalno kompaktne i dobro razdijeljene izborne jedinice, koje se po broju birača međusobno mogu razlikovati najviše 5%. Model se temelji na primjeni klaster analize uz poštovanje zakonom propisanih pravila prema kojima izborne jedinice trebaju imati približno jednak broj birača. Metoda je ilustrirana na primjeru dostup-

[*] Dr. sc. Kristian Sabo, izvanredni profesor Odjela za matematiku, Sveučilište J. J. Strossmayera u Osijeku (associate professor at the Department of Mathematics, University J. J. Strossmayer, Osijek, Croatia, e-mail: ksabo@mathos.hr)

[**] Dr. sc. Rudolf Scitovski, redoviti profesor Odjela za matematiku, Sveučilište J. J. Strossmayera u Osijeku (full professor at the Department of Mathematics, University J. J. Strossmayer, Osijek, Croatia, e-mail: scitowsk@mathos.hr)

[***] Mr. sc. Petar Taler, asistent Odjela za matematiku, Sveučilište J. J. Strossmayera u Osijeku (assistant at the Department of Mathematics, University J. J. Strossmayer, Osijek, Croatia, e-mail: petar@mathos.hr)

nih podataka iz 2007. Tako dobivenu raspodjelu izbornih jedinica ne treba shvatiti kao konačni prijedlog rješenja, već isključivo kao prikaz mogućnosti koje nudi ova metodologija. Prema trenutačno važećem zakonu izbori se u Republici Hrvatskoj provode u 10 izbornih jedinica. Predloženo je nekoliko pristupa poznatih iz literature na osnovi kojih je moguće definirati primjereni broj izbornih jedinica, koje zadržavaju svojstvo maksimalne unutrašnje kompaktnosti i dobre razdijeljenosti.

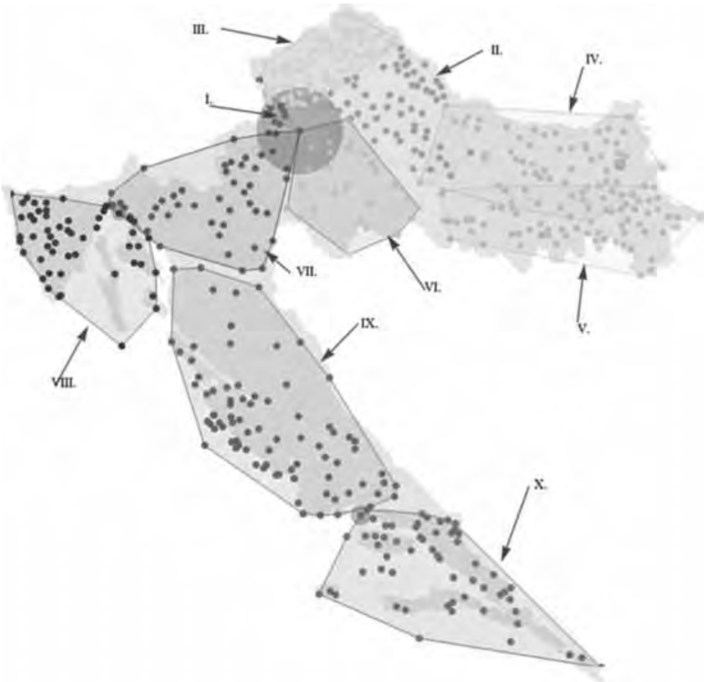
Ključne riječi: grupiranje podataka, klasteri, izborne jedinice

1. Uvod

Prema[1] važećem Zakon o izborima zastupnika u Hrvatski sabor, u Hrvatskoj se primjenjuje razmjerni izborni sustav koji se provodi u 10 velikih višemandatnih izbornih jedinica na koje je podijeljeno područje države. Pritom se u svakoj izbornoj jedinici bira po 14 zastupnika na temelju lista. Čl. 36. Zakona propisuje: »Izborne jedinice određuju se Zakonom o izbornim jedinicama za izbor zastupnika u Hrvatski sabor (Narodne novine, 116/99) tako da se broj birača u izbornim jedinicama ne smije razlikovati više od ±5%. Pri određivanju izbornih jedinica, mora se koliko je to najviše moguće, voditi računa o zakonom utvrđenim područjima županija, gradova i općina u Republici Hrvatskoj.«

Ustavni sud Republike Hrvatske donio je 8. prosinca 2010. Izvješće o nejednakoj težini biračkog glasa u izbornim jedinicama određenima člancima 2. do 11. Zakona o izbornim jedinicama za izbor zastupnika u Zastupnički dom Hrvatskog sabora. U tom se izvješću poziva na podatke Državnog izbornog povjerenstva, prema kojima su se prekomjerna odstupanja u broju birača po pojedinim izbornim jedinicama pojavila već na izborima za zastupnike u Hrvatski sabor održanim 25. studenoga 2007.

[1] Rad je financiran sredstvima projekta Ministarstva znanosti, obrazovanja i športa Republike Hrvatske 235-2352818-1034. Autori zahvaljuju doc. dr. sc. M. Vinkoviću s Pravnog fakulteta Sveučilišta u Osijeku na korisnim primjedbama i savjetima kao i na vrlo konstruktivnoj raspravi. Također, zahvaljuju anonimnom recenzentu koji je svojim primjedbama i sugestijama u znatnoj mjeri pomogao da rad bude bolji.



Slika 1. Podjela na k = 10 izbornih jedinica koja je trenutačno na snazi

Tablica 1. Brojevi birača u pojedinim općim izbornim jedinicama 2007.

Izborna jedinica	Broj birača Q_j	Odstupanje od prosječnog broja birača $(\bar{Q}_j = 382473.1)$ $100\,\frac{Q_j - Q}{\bar{Q}}\,\%$
I.	361.236	-5,55
II.	399.648	4,49
III.	366.005	-4,31
IV.	335.091	-12,39
V.	372.163	-2,70
VI.	356.575	-6,77
VII.	403.812	5,58
VIII.	385.594	0,82
IX.	428.590	12,06
X.	416.017	8,77

U tablici 1 prikazan je broj birača po pojedinoj izbornoj jedinici na tim izborima, a na slici 1 prikazani su najmanji konveksni skupovi koji obuhvaćaju teritorijalne jedinice iz iste izborne jedinice. Na osnovi tih podataka Ustavni sud upozorava: »O ravnomjernoj raspodjeli broja birača po općim izbornim jedinicama (o kojoj raspodjeli neposredno ovisi jednakost težine biračkog glasa) stoga ovisi i zakonitost i opći demokratski karakter cjelokupnih izbora. Štoviše, o tome može ovisiti i ocjena o ustavnosti cjelokupnih izbora: oni bi bili nesuglasni Ustavu ako bi prekomjerno odstupanje u broju birača po pojedinim općim izbornim jedinicama izravno i neposredno utjecalo na izborni rezultat, to jest ako bi dovelo do različitih izbornih rezultata u situaciji kad bi svi ostali elementi izbornoga sustava bili odnosno ostali isti.« Vezano uz određivanje izbornih jedinica, u Izvješću se također navodi: »Područja i granice postojećih upravno-teritorijalnih jedinica (županija, gradova i općina) za tu svrhu nisu u cijelosti prikladni, budući da u njima živi različit broj birača, pa njihovi birački glasovi nemaju jednaku težinu.«

Problem određivanja izbornih jedinica težak je optimizacijski problem, pogotovo u državama s velikim brojem teritorijalnih jedinica. Iz tog je razloga popularan i u znanstveno-stručnoj literaturi gdje se mogu pronaći algoritmi za generiranje raspodjele izbornih jedinica utemeljeni na raznim heurističkim pristupima (Bozkaya i dr., 2003; Chen i Peng, 2008; Grilli, 1999; Ricca i dr., 2008a, 2008b).

U povijesti postoji mnoštvo primjera u kojima su političke stranke pristrano generirale izborne jedinice u skladu sa svojim političkim interesima. Najpoznatiji primjer je iz 1812. kada je država Massachutes pristrano podijeljena u neprirodne izborne jedinice. Prema tadašnjem guverneru Massachutesa Elbridgu Gerryju, takvo prekrajanje izbornih jedinica poznato je kao gerrymandering (v. primjerice Bozkaya i dr., 2003; Grilli, 1999).

Rad namjerava predstaviti jedan matematički model, koji je modificirani oblik modela iz Hess i dr. (1965) na osnovi kojeg je moguće generirati optimalne izborne jedinice koje imaju svojstvo ravnomjerne raspodjele broja birača po tim izbornim jedinicama. Model se temelji na primjeni klaster analize (Kogan, 2007) uz poštovanje zakonom propisanih pravila prema kojima izborne jedinice trebaju imati približno jednak broj birača. U neku izbornu jedinicu ući će one teritorijalne jedinice koje su geografski međusobno bliske. Na taj način izbjegnute su različite subjektivne pretpostavke, kao što je primjerice mogući uspjeh neke političke opcije u toj teritorijalnoj jedinici.

Metoda se ilustrira podacima iz 2007. te tako dobivenu raspodjelu izbornih jedinica ne treba shvatiti kao konačni prijedlog rješenja, već isključivo kao prikaz mogućnosti koje nudi ova metodologija. Problem određivanja izbornih jedinica ovisi o mnoštvu vremenski promjenjivih uvjeta, kao što je primjerice migracija stanovništva. U tom smislu granice izbornih jedinica trebalo bi revidirati u određenim vremenskim intervalima (primjerice, nakon svakog popisa stanovništva). Dakle, model bi trebalo promatrati dinamički, uzimajući u obzir nastale promjene.

Rad je organiziran tako da se u drugom odjeljku definira problem određivanja izbornih jedinica, a u trećem postavlja odgovarajući matematički model te daje algoritam za njegovo rješavanje. Algoritam predstavlja iterativnu proceduru koja polazi od početne aproksimacije. Posebno se razmatra problem odabira dobre početne aproksimacije. Rezultati se ilustriraju u četvrtom odjeljku na primjeru raspodjele broja birača po izbornim jedinicama Republike Hrvatske te se koriste podaci iz 2007. Na kraju tog odjeljka odgovara se na pitanje primjerenog broja izbornih jedinica u Hrvatskoj.

2. Definiranje problema

Općenito se područje neke države promatra organizirano u *m* teritorijalnih jedinica (gradova ili općina), koje su određene geografskim položajem u tzv. Gauss-Krügerovom koordinatnom sustavu točkama $a_i = (x_i, y_i),\ i = 1, \dots, m$. Nadalje pretpostavi se da teritorijalna jedinica a_i ima q_i birača i da je ukupni broj birača Q određen sa $\sum_{i=1}^{m} q_i = Q$. Teritorij promatrane države želi se podijeliti u $k, (1 < k < m)$, izbornih jedinica $\pi_1, \dots, \pi_k$, tako da

(i) izbornu jedinicu sačinjavaju teritorijalne jedinice (gradovi i općine) koje su u nekom razdaljinskom smislu međusobno bliske;

(ii) bude zadovoljen uvjet iz čl. 36/1. Zakona o izborima zastupnika u Hrvatski sabor (NN 116/99), koji glasi: »Izborne jedinice određuju se Zakonom o izbornim jedinicama za izbor zastupnika u Hrvatski sabor tako da se broj birača u izbornim jedinicama ne smije razlikovati više od ±5%.«

Uvjet razdaljinske bliskosti teritorijalnih jedinica neke izborne jedinice može se osigurati primjenom klaster analize uz primjenu odgovarajuće kvazimetričke funkcije $d: \mathbb{R}^2 \times \mathbb{R}^2 \to [0, +\infty\rangle$ (v. npr. Kogan, 2007). Najčešće se u raznim primjenama koristi tzv. *least square* (LS) kvazimetrička funkcija[2] $d(x,y) = \|x-y\|^2$ zbog njezine jednostavnosti i prirodne utemeljenosti. To je razlog zbog kojeg se i ovdje koristi ta kvazimetrička funkcija.

U radu će se postaviti matematički model koji se temelji isključivo na kriterijima (i) i (ii). Radi poboljšanja modela moguće je uzeti u obzir i neke druge kriterije kojima se osiguravaju dodatni uvjeti ravnomjernosti, kao što su primjerice:

- Sociološko-ekonomska homogenost, kojom se osigurava da izborne jedinice imaju približno jednak ukupni dohodak (Bourjolly, 1981; Bozkaya i dr., 2003).
- Sličnost s postojećom konstrukcijom izbornih jedinica, kojom se prilikom revidiranja granica izbornih jedinica osigurava da ne dođe do izrazito velikih promjena prema već postojećim izbornim jedinicama (Bozkaya i dr., 2003).
- Površinska sličnost, kojom se osigurava da izborne jedinice budu približno jednake površine (Bozkaya i dr., 2003).
- Čuvanje većih teritorijalnih jedinica (primjerice županija), kojom se osigurava da se koliko god je moguće granice izbornih jedinica podudaraju s granicama većih teritorijalnih jedinica, kao što su primjerice županije.

Prije postavljanja matematičkog modela ravnomjerne raspodjele broja birača po izbornim jedinicama u sljedećem pododjeljku najprije se analizira čl. 36/1. Zakona.

2.1. Analiza čl. 36. Zakona o izborima zastupnika u Hrvatski sabor

Čl. 36. napisan je nejasno jer otvara mogućnosti različitog tumačenja. Za ilustraciju te tvrdnje pretpostavimo da su $Q_1, \dots, Q_k$ brojevi birača u izbornim jedinicama, pri čemu vrijedi $\sum_{j=1}^{k} Q_j = Q$. Čini se da je članak moguće protumačiti na dva bitno različita načina:

[2] Least squares, najmanji kvadrati. LS-udaljenosti definiraju se kao kvadrati običnih euklidskih udaljenosti. Centar izborne jedinice u tom slučaju predstavljen je točkom (»centroid, težište«) koja ima svojstvo da je suma kvadrata udaljenosti do njezinih teritorijalnih jedinica najmanja.

(A) Za svaki par brojeva birača Q_i i Q_j, $i \neq j$, pri čemu je $Q_i \leq Q_j$ vrijedi

$$100\,\frac{Q_j - Q_i}{Q_i} \leq 5,$$

odnosno bilo koje dvije izborne jedinice smiju se po broju birača razlikovati najviše 5%.

(B) Za broj birača Q_j proizvoljne izborne jedinice π_j vrijedi

$$\left(1 - \frac{5}{100}\right)\frac{Q}{k} \leq Q_j \leq \left(1 + \frac{5}{100}\right)\frac{Q}{k},$$

odnosno broj birača u bilo kojoj izbornoj jedinici Q_j smije se razlikovati od prosječnog broja birača po izbornim jedinicama $\frac{Q}{k}$ najviše 5%.

Nije jasno predviđa li Zakon varijantu (A) ili varijantu (B). Neovisno o tome, iz tablice 1 odmah se vidi da trenutačna podjela Hrvatske na izborne jedinice ne zadovoljava ni uvjet (A) ni uvjet (B). Najveća opća izborna jedinica (IX.) veća je 27,9% od najmanje izborne jedinice (IV.), a isto tako čak šest od deset izbornih jedinica odstupa više od 5% od prosječnog broja birača.

U nastavku se pokazuje da uvjeti (A) i (B) imaju bitno različito značenje.

Ako je zadovoljen uvjet (A), onda je nužno zadovoljen i uvjet (B). Zaista, ako za svaki par brojeva birača Q_i i Q_j, $i \neq j$, pri čemu je $Q_i \leq Q_j$, vrijedi $100\,\frac{Q_j - Q_i}{Q_i} \leq 5$, onda je specijalno

$$100\,\frac{Q_{max} - Q_{min}}{Q_{min}} \leq 5, \tag{1}$$

pri čemu je Q_{min} najmanji, a Q_{max} najveći od brojeva $Q_1, \dots, Q_k$. Sukladno (1), za svaki Q_j vrijedi

$$Q_j \leq Q_{max} \leq \left(1 + \frac{5}{100}\right) Q_{min} \leq \left(1 + \frac{5}{100}\right)\frac{Q}{k}. \tag{2}$$

Analogno za svaki Q_j vrijedi

$$Q_j \geq Q_{min} \geq \frac{Q_{max}}{\left(1 + \frac{5}{100}\right)} \geq \left(1 - \frac{5}{100}\right) Q_{max} \geq \left(1 - \frac{5}{100}\right)\frac{Q}{k}. \tag{3}$$

Iz (2) i (3) slijedi uvjet (B).

Obratno, nije teško vidjeti da uvjet (B) ne povlači uvjet (A), što pokazuje sljedeći jednostavni primjer.

Primjer 1. Neka je $k = 3$ te $Q_1 = 19$, $Q_2 = 20$ i $Q_3 = 21$. Nije teško vidjeti da je uvjet (B) zadovoljen jer vrijedi

$$\left(1-\frac{5}{100}\right)Q_k \leq Q_j \leq \left(1+\frac{5}{100}\right)\frac{Q}{k}, \qquad j = 1{,}2{,}3,$$

dok uvjet (A) nije zadovoljen jer je $100\,\frac{21-19}{19} = \frac{200}{19} \approx 10.53 > 5$.

Općenito vrijedi sljedeća tvrdnja.

Neka je Q ukupni broj birača u državi koja se dijeli na kk izbornih jedinica s odgovarajućim brojem birača $Q_1, \dots, Q_k$ i neka je $q = \frac{100}{41}$. Tada, ako je

$$\left(1-\frac{q}{100}\right)\frac{Q}{k} \leq Q_j \leq \left(1+\frac{q}{100}\right)\frac{Q}{k}, \tag{4}$$

onda za svaki par brojeva birača Q_i, Q_j, $(Q_i \leq Q_j)$ vrijedi

$$100\,\frac{Q_j-Q_i}{Q_i} \leq 100\,\frac{\left(1+\frac{q}{100}\right)\frac{Q}{k}-\left(1-\frac{q}{100}\right)\frac{Q}{k}}{\left(1-\frac{q}{100}\right)\frac{Q}{k}} = \frac{200q}{100-q} = 5.$$

3. Matematički model i algoritam

Pretpostavimo da je područje neke države organizirane u mm teritorijalnih jedinica (gradova ili općina), koje su određene svojim geografskim položajem u tzv. Gauss-Krügerovom koordinatnom sustavu točkama $a_i = (x_i, y_i),\ i = 1, \dots, m$. Nadalje, pretpostavimo da teritorijalna jedinica a_i ima q_i birača i da je ukupni broj birača Q određen sa $\sum_{i=1}^{m} q_i = Q$. Teritorij promatrane države želi se podijeliti u $k, (1 < k < m)$, izbornih jedinica $\pi_1, \dots, \pi_k$ s odgovarajućim brojem birača $Q_1, \dots, Q_k$, poštujući pritom zahtjeve (i), (ii) iz odjeljka 2. Pretpostavimo i da se broj birača Q_j u bilo kojoj izbornoj jedinici smije razlikovati od prosječnog broja birača $\frac{Q}{k}$ po izbornim jedinicama najviše 5%, tj. da vrijedi

$$\left(1-\frac{p}{100}\right)\frac{Q}{k} \leq Q_j \leq \left(1+\frac{p}{100}\right)\frac{Q}{k}.$$

Ako se čl. 36/1. tumači u smislu uvjeta (A), prema raspravi iz odjeljka 2. treba uzeti $p = \frac{100}{41}$, a ako se tumači u smislu uvjeta (B), treba uzeti $p = 5$

Problem određivanja izbornih jedinica izrazito je složen i numerički vrlo zahtjevan matematički problem. Promatrat će se dvije varijante za rješavanje, cjelobrojni pristup odnosno relaksirani linearni pristup.

3.1. Cjelobrojni pristup

Svakoj izbornoj jedinici $\pi_1, \dots, \pi_k$ pridružit će se odgovarajući centri $c_1, \dots, c_k$. Ako se koristi LS-kvazimetrička funkcija, ti centri dobivaju se primjenom težinske aritmetičke sredine (Sabo i dr., 2010). Pritom je suma kvadrata udaljenosti svih teritorijalnih jedinica iz neke izborne jedinice $\pi_j\pi_j$ do njenog centra c_jc_j minimalno moguća. Problem određivanja optimalnih izbornih jedinica svodi se na sljedeći problem klaster analize (Teboulle, 2007):

$$\min_{c_1,\dots,c_k \in \mathbb{R}^2} \sum_{i=1}^{m} \min\{\|a_i - c_1\|^2, \dots, \|a_i - c_k\|^2\}. \tag{5}$$

Općenito svaku teritorijalnu jedinicu $a_i\, a_i$ teorijski i zakonski moguće je u cijelosti ili samo djelomično pridružiti bilo kojoj izbornoj jedinici. Kod *cjelobrojnog pristupa* pretpostavlja se da je svaku teritorijalnu jedinicu a_i a_i moguće jedino u cijelosti pridružiti bilo kojoj izbornoj jedinici. Ako se definira matrica pripadnosti

$$W = (w_{ij}) \epsilon \mathbb{R}^{m\times k}, \quad w_{ij} = \begin{cases} 1, & a_i \in \pi_j, \\ 0, & a_i \notin \pi_j. \end{cases} \tag{6}$$

i ako se vodi računa o zahtjevima (i), (ii) iz odjeljka 2, svi se potrebni uvjeti mogu zapisati na sljedeći način:

$$\sum_{j=1}^{k} w_{ij} = 1, \quad i = 1, \dots, m \tag{7}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{k} w_{ij} q_i = Q \tag{8}$$

$$\left(1 - \frac{p}{100}\right)\frac{Q}{k} \le \sum_{i=1}^{m} w_{ij} q_i \le \left(1 + \frac{p}{100}\right)\frac{Q}{k} \tag{9}$$

$$w_{ij} \in \{0,1\}, \quad i = 1, \dots, m, \quad j = 1, \dots, k \tag{10}$$

Uvjet (7) osigurava da svaka teritorijalna jedinica a_i pripadne točno jednoj izbornoj jedinici π_j, uvjet (8) osigurava da će birači svake teritorijalne jedinice biti uključeni u neku opću izbornu jedinicu, uvjet (9) osigurava ravnomjernost broja birača po općim izbornim jedinicama za najviše p%, a uvjet (10) osigurava da će svaka teritorijalna jedinica $a_i a_i$ u cijelosti pripasti samo jednoj općoj izbornoj jedinici.

Prema Chenu i Pengu (2008), minimizacijska funkcija cilja (5) može se zapisati u obliku

$$\min_{w_{ij}} \sum_{i=1}^{m} \sum_{j=1}^{k} w_{ij} q_i \left\| a_i - \frac{\sum_{s=1}^{m} w_{sj} q_s a_s}{\sum_{s=1}^{m} w_{sj} q_s} \right\|^2 . \tag{11}$$

Ako se pretpostavi da među teritorijalnim jedinicama postoji grad odnosno općina s brojem stanovnika q_{i_0}, koji je više od 5% veći od veličine prosječne izborne jedinice, tj. za čiji broj stanovnika q_{i_0} vrijedi

$$q_{i_0} \geq \left(1 + \frac{p}{100}\right) \frac{Q}{k}, \tag{12}$$

optimizacijski problem (11) uz uvjete (7) – (10) neće imati rješenje. Takva situacija događa se u Hrvatskoj za Grad Zagreb u slučaju 10 izbornih jedinica. U tom slučaju mora se omogućiti da se takva teritorijalna jedinica razdijeli u više izbornih jedinica.

U matematičkom modelu to znači da teorijski treba dopustiti diobu te teritorijalne jedinice na svih k izbornih jedinica. To će se postići tako da se uvede skup indeksa I_0 onih teritorijalnih jedinica za koje vrijedi (12) te se umjesto uvjeta (10) uvodi sljedeći uvjet

$$w_{ij} = \begin{cases} [0,1], & i \in I_0 \\ \{0,1\}, & i \in \{1, \dots, m\} \setminus I_0 \end{cases}, \quad j = 1, \dots, k \tag{13}$$

Optimizacijski problem (11) uz uvjete (7) – (9) te (13) problem je nelinearne globalne optimizacije s ograničenjima (Floudas i Gounaris, 2009; Neumaier, 2004) s iznimno velikim brojem varijabli i mnogo potencijalnih lokalnih rješenja. Primjerice, za problem određivanja 10 optimalnih izbornih jedinica u Hrvatskoj pojavljuje se 5.560 varijabli. Direktno rješavanje tog problema s numeričkog aspekta izrazito je zahtjevno.

Umjesto rješavanja tog problema globalne optimizacije, temeljem rada Ng (2000) konstruirat će se algoritam koji uz dobru početnu aproksimaciju daje rješenje za koje se s velikom sigurnošću može tvrditi da je optimalno.[3]

Algoritam 1.

Korak 0: Učitati
$1 \leq k \leq m;\ I = \{1, \ldots, m\};\ J = \{1, \ldots, k\};\ \ A = \{a_i \in \mathbb{R}^2 : i \in I\}$.
Izabrati međusobno različite centre $c_1, \ldots, c_k$ koji predstavljaju početnu aproksimaciju centara izbornih jedinica;

Korak 1: Riješiti problem cjelobrojnog programiranja (Sierksma, 2002)

$$\min_{w_{ij}} \sum_{i=1}^{m} \sum_{j=1}^{k} w_{ij} q_i \left\| a_i - c_j \right\|^2,$$

uz uvjete (7) – (9) te (13). Rješenje je matrica koju se označi s $\hat{W} = \left(\hat{w}_{ij}\right)$;

Korak 2: Izračunati nove centre kao težinske aritmetičke sredine

$$\hat{c}_j := \frac{\sum_{s=1}^{m} \hat{w}_{sj} q_s a_s}{\sum_{s=1}^{m} \hat{w}_{sj} q_s}, \ \ j = 1, \ldots, k$$

Korak 3: Ako postoji $j \in J$, takav da je $c_j \neq \hat{c}_j$, staviti $c_j = \hat{c}_j$ za svaki $j \in J$ i prijeći na *korak 1*;

Inače staviti $W^* = \hat{W}, c_j^* = \hat{c}_j$ za svaki $j \in J$ i zaustaviti algoritam.

Kao što se vidi, algoritam 1 konstruiran je tako da sukcesivno rješava optimizacijski problem (5) s poznatim centrima uz uvjete (7) – (9) te (13) (korak 1) i određuje nove centre klastera kao težinske aritmetičke sredine podataka a_i s težinama dobivenim u koraku 1 (korak 2). Uz pretpostavku da se algoritam pokrenulo s dobrom početnom aproksimacijom centara $c_1, \ldots, c_k$, s velikom sigurnošću može se tvrditi da se kao rezultat izvođenja algoritma dobilo optimalne centre $c_1^*, \ldots, c_k^*$ izbornih jedinica, a same izborne jedinice definirane su tada sa

$$\pi_j^* = \left\{a_i \in \mathcal{A} : w_{ij}^* \neq 0\right\}, \ \ j = 1, \ldots, k$$

[3] Provedba algoritma 1, što uključuje određivanje početne aproksimacije na primjeru podataka za Hrvatsku s 10 izbornih jedinica na računalu s Intel Core i5 760 procesorom i 8GB radne memorije traje 5 minuta. U svrhu rješavanja linearnog odnosno cjelobrojnog programiranja koriste se rutine iz GNU Linear Programming Kit (GLPK) biblioteke (http://www.gnu.org/s/glpk).

Pritom treba imati na umu da za $a_i \in \pi_j^*$ mogu nastupiti dvije mogućnosti:

- ako je $w_{ij}^* = 1$, teritorijalna jedinica a_i cijela je sadržana u izbornoj jedinici π_j^*;
- ako je w_{ij}^* broj između 0 i 1, izborna jedinica π_j^* sadržava w_{ij}^*-ti dio teritorijalne jedinice a_i (ta mogućnost dogodit će se jedino onda ako je $i \in I_0$, gdje je I_0 skup indeksa onih teritorijalnih jedinica za koje vrijedi (12)).

Metoda za određivanje dobre početne aproksimacije opisuje se u sljedećem pododjeljku.

3.2. Relaksirani linearni pristup

Uvjet (10) garantira da granice dobivenih izbornih jedinica neće dijeliti birače nijedne općine odnosno grada. Međutim sukladno izvješću Ustavnog suda, »područja i granice postojećih upravno-teritorijalnih jedinica (županija, gradova i općina) za tu svrhu nisu u cijelosti prikladni, budući da u njima živi različit broj birača, pa njihovi birački glasovi nemaju jednaku težinu«. Zbog toga je ovaj uvjet dopušteno relaksirati tako da granice izbornih jedinica smiju dijeliti neke gradove ili općine. To praktično znači da se nekoj teritorijalnoj jedinici a_i neće pridružiti broj 0 ili 1 prema (6), već niz realnih brojeva iz intervala $[0, 1]$

$$w_{ij} \in [0,1], \quad j = 1, \ldots, k \tag{14}$$

za koje također mora vrijediti uvjet (7):

$$\sum_{j=1}^{k} w_{ij} = 1, \quad i = 1, \ldots, m.$$

To znači da se birači iz teritorijalne jedinice a_i mogu podijeliti po različitim izbornim jedinicama poštujući pritom druge uvjete iz modela. Naravno, u tom slučaju uvjet (9) potrebno je redefinirati u sljedeći oblik

$$\left\lfloor \left(1 - \frac{p}{100}\right)\frac{Q}{k} \right\rfloor + 1 \le \sum_{i=1}^{m} w_{ij} q_i \le \left\lfloor \left(1 + \frac{p}{100}\right)\frac{Q}{k} \right\rfloor - 1 \tag{15}$$

kako bi se na taj način osigurao cjelobrojni broj birača u pojedinim izbornim jedinicama.

Uz takvu relaksaciju problem globalne optimizacije (5) uz uvjete (7), (8), (14) i (15) postaje nešto jednostavniji problem od originalnog. No, i u tom slučaju umjesto globalne optimizacije za njegovo rješavanje primijenit će se modifikacija algoritma 1, utemeljena na pristupu iz Ng (2000). Za razliku od originalnog cjelobrojnog problema iz koraka 1, u tom slučaju potrebno je rješavati znatno jednostavniji problem linearnog programiranja (Sierksma, 2002). Zato umjesto traženja dobre početne aproksimacije, sukladno Leischu (2006), iterativni se postupak pokreće primjerice s 1.000 različitih slučajno generiranih početnih centara te se kao rješenje uzima ono koje daje najmanju vrijednost funkcije cilja (5).

Rješenje dobiveno tim relaksiranim linearnim pristupom može se uzeti kao konačno rješenje (uz napomenu da će se u tom slučaju eventualno dogoditi da granice izbornih jedinica, osim velikih gradova, dijele i neke druge gradove ili općine) ili kao dobra početna aproksimaciju kod cjelobrojnog pristupa.

4. Problem definiranja izbornih jedinica u Hrvatskoj

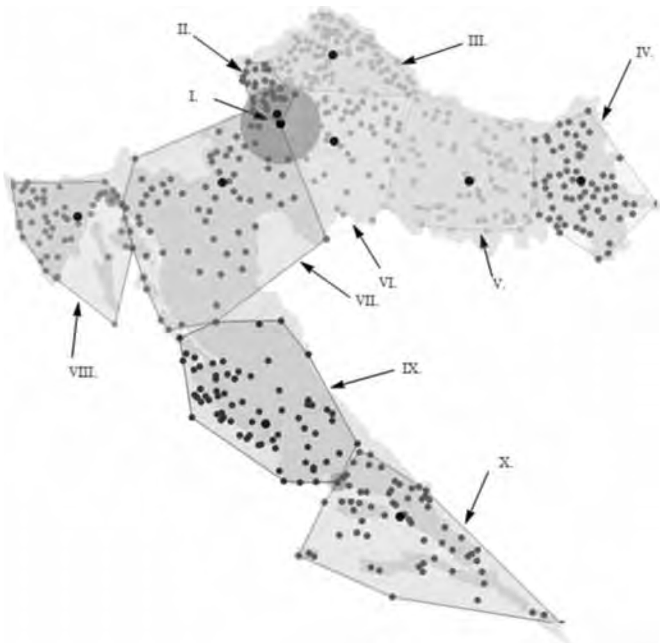
Područje Hrvatske organizirano je u $m = 556$ teritorijalnih jedinica (gradova ili općina), koje su određene svojim geografskim položajem u tzv. Gauss-Krügerovom koordinatnom sustavu točkama $a_i = (x_i, y_i),\ i = 1, \dots, m$. Broj birača po teritorijalnim jedinicama iz 2007. označen je sa q_i, a ukupni broj birača Q određen je sa $\sum_{i=1}^{m} q_i = Q$. Čitav teritorij želi se podijeliti u $k, (1 < k < m)$ izbornih jedinica $\pi_1, \dots, \pi_k$ s odgovarajućim brojem birača $Q_1, \dots, Q_k$, poštujući pritom zahtjeve (i), (ii) iz odjeljka 2. Broj birača po općim izbornim jedinicama ne smije se razlikovati više od 5%, tj. mora vrijediti zahtjev (9) uz $p = \frac{100}{41}$. Najprije će se potražiti rješenje tog problema za $k = 10$ uz primjenu relaksiranog linearnog pristupa i cjelobrojnog pristupa. Dobiveni rezultati usporedit će se s konfiguracijom izbornih jedinica pred izbore 2007. Nakon toga provest će se analiza optimalnog broja izbornih jedinica, za koji se dopušta da se kreće od $k = 2, \dots, 10$, te na taj način predložiti primjereni broj izbornih jedinica.

4.1. Primjena relaksiranog linearnog pristupa na model s 10 izbornih jedinica

Primjenom tog pristupa dobiva se 10 izbornih jedinica s brojevima birača koji su prikazani u tablici 2. Na slici 2 različitim bojama označeni su najmanji konveksni skupovi koji sadržavaju one teritorijalne jedinice koje se nalaze u istim izbornim jedinicama. Crnim točkama označena su geografska mjesta centara pojedinih klastera (izbornih jedinica). Iz tablice se vidi da se nijedna izborna jedinica ne razlikuje od prosjeka više od $p = \frac{100}{41}\%$, a isto tako da se najveća od najmanje izborne jedinice razlikuje manje od 5%. U skladu s očekivanjima, sljedeće teritorijalne jedinice podijeljene su između izbornih jedinica: Dvor na Uni (u VI. izbornoj jedinici 0,44%, a u VII. izbornoj jedinici 99,56%), Gospić (u VII. izbornoj jedinici 65,38%, a u IX. izbornoj jedinici 34,62%), Koška (u IV. izbornoj jedinici 45,5%, a u V. izbornoj jedinici 55,5%), Split (u IX. izbornoj jedinici 41,5%, a u X. izbornoj jedinici 58,5%) Zagreb (u I. izbornoj jedinici 56,3%, II. izbornoj jedinici 41,5%, VII. izbornoj jedinici 2,2%), Punat (u VII. izbornoj jedinici 40%, a u VIII. izbornoj jedinici 60%) i Podravske Sesvete (u III. izbornoj jedinici 39,8%, a u V. izbornoj jedinici 50,2%), Kloštar Podravski (u V. izbornoj jedinici 32,1%, a u III. izbornoj jedinici 67,9%). To rješenje može se uzeti kao rješenje problema ili kao dobra početna aproksimacija na cjelobrojni pristup.

Tablica 2. Brojevi birača u pojedinim izbornim jedinicama koji su dobiveni primjenom relaksiranog linearnog pristupa

Izborna jedinica	Broj birača Q_j	Odstupanje od prosječnog broja birača ($\bar{Q}_j = 382473.1$) $100\,\frac{Q_j - Q}{\bar{Q}}\,\%$
I.	391.801	2,44
II.	391.801	2,44
III.	373.146	-2,44
IV.	373.145	-2,44
V.	373.145	-2,44
VI.	373.145	-2,44
VII.	391.801	-2,44
VIII.	391.801	2,44
IX.	391.801	2,44
X.	391.801	2,44



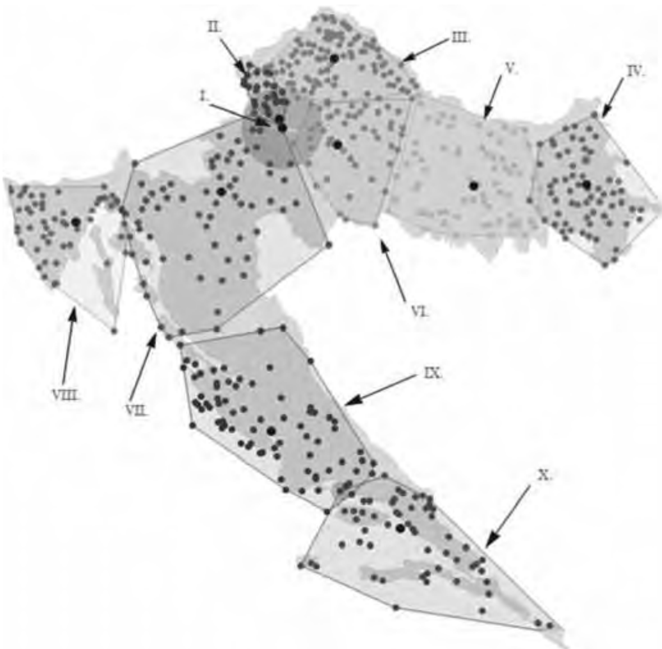
Slika 2. Podjela na k=10 izbornih jedinica prema relaksiranom linearnom pristupu

4.2. Primjena cjelobrojnog pristupa na model s 10 izbornih jedinica

Primjenom algoritma 1, uz početnu aproksimaciju dobivenu relaksiranim linearnim pristupom iz pododjeljka 4.1., dobiva se 10 izbornih jedinica čiji su brojevi stanovnika prikazani u tablici 3. Primjećuje se da je uvjet ravnomjernosti broja birača po općim izbornim jedinicama zadovoljen.

Tablica 3. Brojevi birača u pojedinim izbornim jedinicama koji su dobiveni primjenom cjelobrojnog pristupa

Izborna jedinica	Broj birača Q_j	Odstupanje od prosječnog broja birača ($\bar{Q}_j = 382473.1$) $100\,\frac{Q_j - Q}{\bar{Q}}\,\%$
I.	391.801	2,44
II.	391.415	2,34
III.	380.918	-0,40
IV.	373.619	-2,31
V.	373.238	-2,41
VI.	373.762	-2,28
VII.	373.145	-2,44
VIII.	391.000	2,23
IX.	384.194	0,45
X.	391.639	2,40



Slika 3. Podjela na k=10 izbornih jedinica prema cjelobrojnom pristupu

Ovdje svaka teritorijalna jedinica pripada točno jednoj izbornoj jedinici, osim grada Zagreba, koji je potrebno podijeliti između I. (56,3%), II. (42,7%) i VII. (1%) izborne jedinice. Određivanje 10 izbornih jedinica dovelo je do neprirodne situacije prema kojoj je grad Zagreb podijeljen na tri znatno nejednaka dijela (slika 3), pri čemu je onaj najmanji (1%) pripao VII. izbornoj jedinici. Treba primijetiti da ostaje problem načina podjele neke teritorijalne izborne jedinice na više izbornih jedinica.

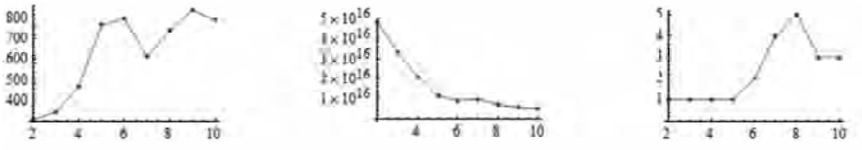
4.3. Optimizacija broja izbornih jedinica

Prirodno se postavlja pitanje koliki bi broj izbornih jedinica bio najprimjereniji sadašnjem geopolitičkom statusu Hrvatske. Drugim riječima, koliko bi trebalo biti izbornih jedinica i kakva bi trebala biti njihova konfiguracija tako da bude osigurana kompaktnost interesa, pri čemu treba zadržati ravnomjernost broja birača po općim izbornim jedinicama kao što je predviđeno matematičkim modelom?

Prevedeno na jezik matematičkog modela, pitanje se svodi na određivanje optimalne particije s brojem klastera, koji osigurava dobru kompaktnost i razdvojenost pojedinih klastera u particiji. Na to pitanje u literaturi se pokušava odgovoriti primjenom različitih indeksa, kao što su: Davies-Bouldinov indeks, Calinski-Harabaszeov indeks, brzina pada vrijednosti minimizirajuće funkcije (Kogan, 2007; Sabo i dr., 2010; Gan i dr., 2007). Isto tako smisleno je minimizirati broj izbornih jedinica na koje je potrebno podijeliti najveću teritorijalnu jedinicu – grad Zagreb. Obično je potrebno kombinirati sve indekse te za broj izbornih jedinica odabrati onaj na koji upućuje najveći broj indeksa.

Veća vrijednost Calinski-Harabaszovog indeksa odgovara primjerenijem broju izbornih jedinica. Na slici 4.a vidi se da bi se sukladno tom kriteriju za broj izbornih jedinica trebalo uzeti 5, 6, 9 ili 10. Promatra li se vrijednost minimizirajuće funkcije (slika 4.b), vidi se da se njezina brzina pada znatno smanjuje na 5 odnosno 6 izbornih jedinica, što znači da prema tom kriteriju primjeren broj izbornih jedinica iznosi 5 ili 6. Isto tako, s 5 odnosno 6 izbornih jedinica (slika 4.c) najviše će se smanjiti dijeljenje grada Zagreba. Davies-Bouldinov indeks ne daje značajne razlike u broju izbornih jedinica.

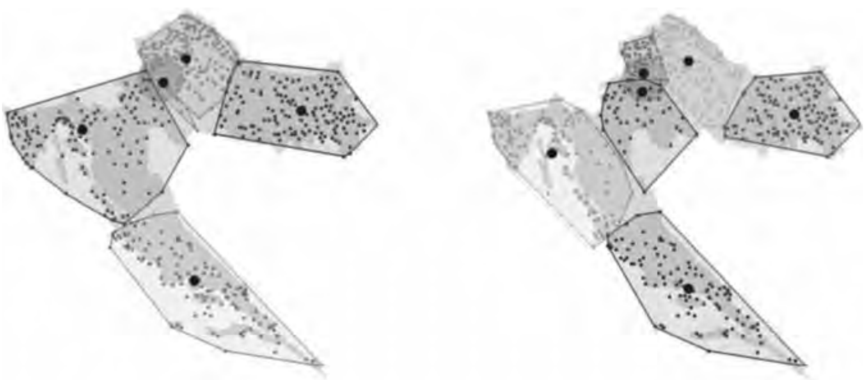
a) Calinski-Harabaszov indeks u ovisnosti o broju izbornih jedinica	b) Vrijednost minimizirajuće funkcije u ovisnosti o broju izbornih jedinica	c) Broj izbornih jedinica u koje je je razdijeljen grad Zagreb u ovisnosti o o ukupnom broju izbornih jedinica



Slika 4. Vrijednosti različitih indeksa u ovisnosti o broju izbornih jedinica

Primjena navedenih indeksa na problem određivanja optimalnog broja izbornih jedinica u Republici Hrvatskoj upućuje na to da bi primjereni broj izbornih jedinica trebao biti 5, u slučaju da se ne želi dijeliti grad Zagreb, odnosno 6 ako se dopusti dijeljenje grada Zagreba. Na slici 5 prikazana je podjela Hrvatske na 5 odnosno 6 izbornih jedinica. Pritom su crnim točkama označena geografska mjesta centara pojedinih klastera (izbornih jedinica).

$k = 5k = 5$ izbornih jedinica $\qquad$ $k = 6k = 6$ izbornih jedinica



Slika 5. Podjela na k=5 i k=6 izbornih jedinica prema cjelobrojnom pristupu

5. Zaključak

U radu je predložen matematički model temeljem kojeg je moguće definirati prostornu konfiguraciju i broj birača teritorijalnih izbornih jedinica. Prema trenutačno važećem zakonu, izbori se u Hrvatskoj provode u 10 izbornih jedinica. Predloženim modelom predviđena je i mogućnost određivanja broja izbornih jedinica kojim se osigurava maksimalna unutrašnja kompaktnost i dobra razdijeljenost tih izbornih jedinica. Također, prilikom pisanja zakona trebalo bi izbjegavati rečenične konstrukcije koje otvaraju mogućnosti različitog tumačenja, kao što je primjerice čl. 36. Zakona o izborima zastupnika u Hrvatski sabor.

U okviru daljnjih istraživanja moguće je predloženi matematički model poopćiti tako da se osim teritorijalne gustoće stanovnika uzmu u obzir i neki drugi važni kriteriji, kao što su primjerice prometna, kulturna i ekonomska povezanost područja. Uz takve pretpostavke posebno važan otvoreni problem također je i određivanje primjerenog broja izbornih jedinica.

Ako matematički model predviđa mogućnost da je neka teritorijalna jedinica podijeljena u više izbornih jedinica, kao što je primjerice opisani relaksirani linearni pristup, otvara se nov optimizacijski problem načina podjele te teritorijalne jedinice.

Također, smisleno je problem određivanja izbornih jedinica promatrati i tako da se područje Hrvatske najprije primjenom klaster analize podijeli u nekoliko geografski, kulturno, prometno i ekonomski bliskih područja, neovisno o broju glasača, te da svako tako dobiveno područje bude jedna izborna jedinica. Pritom se broj zastupnika koji se biraju iz promatrane izborne jedinice određuje tako da bude razmjeran broju glasača te izborne jedinice. Uz takvu formulaciju otvara se nov matematički problem definiranja broja zastupničkih mjesta za svaku izbornu jedinicu tako da ukupni broj zastupnika po svim izbornim jedinicama bude unaprijed zadan cijeli broj. I u tom slučaju ostaje otvoren problem definiranja primjerenog broja izbornih jedinica.

Literatura

Bourjolly, J. M., G. Laporte, J. M. Rousseau (1981) Découpage électoral automatisé à I'le de Montréal. INFOR, 19: 113–124

Bozkaya, B., E. Erkut, G. Laporte (2003) A tabu search heuristic and adaptive memory procedure for political districting. Europian Journal of Operational Reacearch, 12–26

Chen, H., J. Peng (2008) 0–1 semidefinite programming for graph-cut clustering: modelling and approximation. In: P. M. Pardalos, P. Hansen (eds.) Data Mining and Mathematical Programming, 15–39

Floudas, C. A., C. E. Gounaris (2009) A review of recent advances in global optimization. Journal of Global Optimization, 45: 3–38

Gan, G., C. Ma, J. Wu (2007) Data Clustering: Theory, Algorithms, and Applications. Philadelphia: SIAM

Grilli de Cortona, P., C. Manzi, A. Pennisi, F. Ricca, B. Simeone (1999) Evaluation and optimization of electoral systems. Philadelphia: SIAM Monographs on Discrete Mathematics

Hess, S. W., J. B. Weaver, H. J. Whelan, P. A. Zitlau (1965) Nonpartisian political redistricting by computer. Operations Reasearch, 13: 998–1006

Izvješće (2010) Izvješće o nejednakoj težini biračkog glasa u izbornim jedinicama određenima člancima 2. do 11. Zakona o izbornim jedinicama za izbor zastupnika u Zastupnički dom Hrvatskog sabora (NN 116/99), NN 142/10

Kogan, J. (2007) Introduction to clustering large and high-dimensional dana. Cambridge University Press

Leisch, F. (2006) A toolbox for K-centroids cluster analysis. Computational Statistics & Data Analysis 51: 526–544

Neumaier, A. (2006) Complete search in continuous global ooptimization and constraint satisfaction. Acta Numerica, 271–369

Ng, M. K. (2000) A note on constrained k-means algorithms. Pattern Recognition 33: 525–519

Ricca, F., B. Simeoni (2008) Local search algorithms for political districting. European Journal of Operational Reacearch, 1409–1426

Ricca, F., A. Scozzari, B. Simeoni (2008) Weighted Voronoi region algorithms for political districting. Mathematical and Computer Modelling, 48: 1468–1477

Sabo, K., R. Scitovski, I. Vazler (2010) Grupiranje podataka: klasteri. Osječki matematički list 10: 149–176

Sabo, K., R. Scitovski, I. Vazler, M. Zekić-Sušac (2011) Mathematical models of natural gas consumption. Energy Conversion and Management, 52: 1721–1727

Sabo, K., R. Scitovski, I. Vazler: One-dimensional center-based l_1l_1-clustering method. Optimization Letters (prihvaćno za objavljivanje)

Sierksma, G. (2002) Linear and Integer Programming. Theory and Practice}, 2nd edition. New York: Marcel Dekker

Teboulle, M. (2007) A unified continuous optimization framework for center-based clustering methods. Journal of Machine Learning Research, 8: 65–102

Zakon o izborima zastupnika u Hrvatski sabor, NN 116/99, 109/00, 53/03, 69/03 – pročišćeni tekst, 167/03, 19/07, 20/09, 145/10

Zakon o izbornim jedinicama za izbor zastupnika u Hrvatski sabor, NN 116/99

UNIFORM DISTRIBUTION OF THE NUMBER OF VOTERS PER CONSTITUENCY ON THE BASIS OF A MATHEMATICAL MODEL

Summary

The paper presents a mathematical model on the basis of which it is possible to define maximum compact well-separated constituencies, which can vary with respect to the number of voters by 5 per cent at most. The model is set up so as not to favor any political option and it is based on the application of cluster analysis, taking into account the rule according to which constituencies should have roughly the same number of voters. The method is illustrated on the example of the available data from 2007, so that the distribution of constituencies obtained in this way should not be taken as the final solution proposal, but only as a demonstration of the possibilities offered by this methodology. According to the current law, the elections in Croatia are carried out in 10 constituencies. In this paper, several approaches known from the literature are proposed on the basis of which it is possible to determine the appropriate number of constituencies, which retain the property of maximum internal compactness and satisfactory well-separation.

Key words: *data clustering, clustering, constituencies*